



Multiscale top-hat selection transform based infrared and visual image fusion with emphasis on extracting regions of interest



Xiangzhi Bai^{a,b,c,*}, Xiaowu Chen^b, Fugen Zhou^{a,c}, Zhaoying Liu^a, Bindang Xue^a

^a Image Processing Center, Beijing University of Aeronautics and Astronautics, 100191 Beijing, China

^b State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China

^c Beijing Key Laboratory of Digital Media, Beihang University, Beijing 100191, China

HIGHLIGHTS

- A multiscale top-hat selection transform based algorithm is proposed.
- Top-hat selection transform well differentiates and extracts the regions of interest.
- The algorithm appropriately combines the extracted useful image information. • The algorithm produces a clear result contained more useful image information.

ARTICLE INFO

Article history:

Received 29 September 2012

Available online 19 March 2013

Keywords:

Infrared and visual image fusion

Top-hat selection transform

Multiscale morphology

Region extraction

Mathematical morphology

ABSTRACT

To effectively combine regions of interest in original infrared and visual images, an adaptively weighted infrared and visual image fusion algorithm is developed based on the multiscale top-hat selection transform. First, the multiscale top-hat selection transform using multiscale structuring elements with increasing sizes is discussed. Second, the image regions of the original infrared and visual images at each scale are extracted by using the multiscale top-hat selection transform. Third, the final fusion regions are constructed from the extracted multiscale image regions. Finally, the final fusion regions are combined into a base image calculated from the original images to form the final fusion result. The combination of the final fusion regions uses the adaptive weight strategy, and the weights are adaptively obtained based on the importance of the extracted features. In the paper, we compare seven image fusion methods: wavelet pyramid algorithm (WP), shift invariant discrete wavelet transform algorithm (SIDWT), Laplacian pyramid algorithm (LP), morphological pyramid algorithm (MP), multiscale morphology based algorithm (MSM), center-surround top-hat transform based algorithm (CSTHT), and the proposed multiscale top-hat selection transform based algorithm. These seven methods are compared over five different publicly available image sets using three metrics of spatial frequency, mean gradient, and Q . The results show that the proposed algorithm is effective and may be useful for the applications related to the infrared and visual image fusion.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Fusion of image information from multiple imaging sensors is an important technique in different applications [1–4]. Especially, fusion of images obtained from infrared and visual imaging sensors is important to improve the performance of some military or civil applications, such as image based guidance, security surveillance and targeting [3,4]. Infrared (IR) images contain important image regions which could not be displayed by the visual image. However, a visual image usually contains more image details than an infrared image. So, to obtain an image which contains the important image regions and more image details, the technique of infra-

red and visual image fusion is an important way to achieve this purpose. Usually, the used original infrared and visual images should be already registered. To obtain a good fusion result image with more useful image information, different algorithms have been proposed. Direct average algorithm is widely used in different applications [5,6]. But, this algorithm smoothes some important regions in the infrared image and image details in the visual image, which produces an un-clear result. The wavelet pyramid (WP) and curvelet transforms decompose the original images into different images representing the multiscale features of the original image, and these multiscale features are used for image fusion [6–13]. But, some important image details may also be smoothed, which will affect the performance of these algorithms. Although the shift invariant discrete wavelet transform algorithm (SIDWT) [10,11], which is the improved wavelet pyramid algorithm, performs better

* Corresponding author at: Image Processing Center, Beijing University of Aeronautics and Astronautics, 100191 Beijing, China. Tel.: +86 10 82338578.

E-mail address: jackybxz@buaa.edu.cn (X. Bai).

for detail protection, some image details may be still smoothed and the result image may be not clear. Laplacian pyramid algorithm (LP) [6,11,14,15] could also extract the multiscale image features for image fusion. Although the result image is clear because some edge features may be strengthened, some image details may be smoothed and the contrast of the result image may be not good. Independent component analysis (ICA) or principal component analysis (PCA) based algorithms extract the main information of the original images and combine the extracted information together to form the final fusion result [16–18]. However, some detailed information which represents important image regions may not be preserved in the final fusion result. Neural networks are also applied for image fusion, which are mainly used for the multi-focus image fusion [19,20].

Mathematical morphology based algorithms are effective ways for image fusion [5,21–25]. These algorithms use morphological operations to extract the useful image features and combine these image features together to form the final fusion image. Among these morphological operations used for feature extraction, top-hat transforms show promising results. Through the multiscale extension of the classical top-hat transform by using the multiscale morphological theory (MSM), image features in the original images may be well extracted and combined into the final fusion image [5]. However, because the performance of opening and closing in the classical top-hat transform may be not good for the extraction of the regions of interest [22], the image details in the fusion result are not clear. To improve the performance of top-hat transform through constructing structuring elements, the center-surround top-hat transform based algorithm (CSTHT) is proposed [23]. Although it is effective for obtaining a fusion result with good contrast [23], some image details are smoothed. This may make the final result un-clear. Also, extracting the focus regions from the multiscale features by top-hat transform performs well for multi-focus image fusion [25]. The algorithm may be not effective for infrared and visual image fusion. Moreover, by utilizing the morphological operations in the pyramid decomposition theory, the constructed morphological pyramid algorithm (MP) could also extract the multiscale pyramid features for image fusion [21]. However, because of the sampling, the morphological pyramid algorithm may produce some artifacts in the result image. This will affect the application of the fusion result.

Actually, the regions of interest in infrared images are usually bright or dim regions comparing with the surrounding regions. Also, the useful image details in visual image are image regions which are different from the surrounding regions. These regions of interest in the infrared and visual images usually represent the useful and important image regions which are different from the surrounding regions. And, these regions of interest are the useful and important image information which should be combined into the final fusion image. So, one important part of infrared and visual image fusion is effectively extracting the important regions of interest which are different from the surrounding regions. The top-hat selection transform [24], which is the modification of the classical top-hat transform, could selectively output image regions following different application purposes. So, top-hat selection transform may be well used for the extraction of the regions of interest in infrared and visual images. Then, based on top-hat selection transform, an effective infrared and visual image fusion algorithm may be constructed.

Through extracting the regions of interest, a multiscale top-hat selection transform based infrared and visual image fusion algorithm is proposed in this paper. First, specifying the top-hat selection transform to extract the regions of interest and constructing multiscale top-hat selection transform using multiscale structuring elements with increasing sizes are discussed. Second, the multiscale image regions of the original infrared and visual images are

extracted. Third, the regions of interest represented by the final fusion regions are constructed from the extracted multiscale image regions. Finally, the final fusion image is obtained through importing the extracted regions of interest into a base image using the weight strategy. Because of the effectively extracting the regions of interest using the specified top-hat selection transform, the proposed algorithm performs well for infrared and visual image fusion. Experimental results on infrared and visual images verified the good performance of the proposed algorithm.

2. Top-hat selection transform

2.1. Mathematical morphology

Mathematical morphology has been an important theory for image processing and pattern recognition after being proposed [21]. Morphological operations are mainly based on set theory. And, the two basic operations are dilation and erosion which use two sets. One set is the image and the other set is called structuring element.

Let f and B represent the grayscale image and structuring element, respectively. The dilation and erosion of $f(x, y)$ by $B(u, v)$, represented by $f \oplus B$ and $f \ominus B$, are given as follows,

$$f \oplus B = \max_{u,v} (f(x-u, y-v) + B(u, v)) \quad (1)$$

$$f \ominus B = \min_{u,v} (f(x+u, y+v) - B(u, v)) \quad (2)$$

(x, y) and (u, v) are the pixel coordinates of f and B , respectively.

Through combining the dilation and erosion, the opening and closing of $f(x, y)$ by $B(u, v)$, represented by $f \circ B$ and $f \bullet B$, are as follows,

$$f \circ B = (f \ominus B) \oplus B, \quad (3)$$

$$f \bullet B = (f \oplus B) \ominus B. \quad (4)$$

By comparing the result of opening or closing with the original image, the white top-hat transform and black top-hat transform of image f , represented by WTH and BTH , are as follows,

$$WTH(x, y) = f(x, y) - f \circ B(x, y), \quad (5)$$

$$BTH(x, y) = f \bullet B(x, y) - f(x, y). \quad (6)$$

Opening is usually used to smooth bright image regions whereas closing is usually used to smooth dim image regions. So, WTH is used to extract bright image regions and BTH is used to extract dim image regions.

2.2. Top-hat selection transform

Opening smoothes bright image regions and closing smoothes dim image regions. So, $f(x, y) \geq f \circ B(x, y)$, $f(x, y) \leq f \bullet B(x, y)$. Take WTH as an example, because $f(x, y) \geq f \circ B(x, y)$, WTH could be recognized as a selective output, which could be rewritten as follows [24].

$$WTH(x, y) = \begin{cases} f(x, y) - f \circ B(x, y), & \text{if } f(x, y) - f \circ B(x, y) \geq 0 \\ k, & \text{else,} \end{cases} \quad (7)$$

where k is an arbitrary value.

Because top-hat transform does not well utilize the difference information between the regions of interest and the surrounding regions, the performance of top-hat transform is not effective. Through expanding and controlling the difference information between the original image and the result of opening or closing following expression (7), the performance of top-hat transform

could be improved. Based on this idea, the top-hat selection transform was proposed [24].

The white top-hat selection transform is given below.

$$WTHS(x, y) = \begin{cases} f(x, y) - f \circ B(x, y), & \text{if } t_1 \leq f(x, y) - f \circ B(x, y) \leq t_2 \\ t, & \text{else} \end{cases} \quad (8)$$

where $t \in (-\infty, +\infty)$, $t_1 \in [0, +\infty)$ and $t_2 \in [0, +\infty)$.

Also, the black top-hat selection transform is given below.

$$BTHS(x, y) = \begin{cases} f \bullet B(x, y) - f(x, y), & \text{if } t_1 \leq f \bullet B(x, y) - f(x, y) \leq t_2 \\ t, & \text{else} \end{cases} \quad (9)$$

$WTHS$ and $BTHS$ indicate that, through setting the values of t , t_1 and t_2 in purpose, the difference information between the original image and the result of opening or closing could be well used for different applications [24], which improves the performance of top-hat transform.

A flow graph of the calculation of $WTHS$ and $BTHS$ is shown in Fig. 1, which clearly shows the calculation of top-hat selection transforms from the basic morphological operations. The basic morphological dilation and erosion produce the opening and closing operations. And, the difference operations of the original image and the result of opening or closing produce the classical top-hat transforms. By importing t , t_1 and t_2 into the classical top-hat transforms, the top-hat selection transforms are obtained.

3. Algorithm

3.1. Extracting multiscale image regions

3.1.1. Extracting regions of interest using top-hat selection transform

In infrared image, the regions of interest are usually bright or dim regions whose main features are larger or smaller gray values than the surrounding regions. So, after the opening or closing, the changing of the gray values in the regions of interest is usually larger than the surrounding regions. Thus, it is easy to extract the regions of interest through specifying the top-hat selection transform as follows [24].

$$WTHS(x, y) = \begin{cases} f(x, y) - f \circ B(x, y), & \text{if } f(x, y) - f \circ B(x, y) \geq nL \\ 0, & \text{else} \end{cases} \quad (10)$$

$$BTHS(x, y) = \begin{cases} f \bullet B(x, y) - f(x, y), & \text{if } f \bullet B(x, y) - f(x, y) \geq nL \\ 0, & \text{else} \end{cases} \quad (11)$$

This definition indicates that, only the regions whose gray values are larger than the surrounding region at least nL are the real regions of interest. This property is useful for extracting the regions of interest in infrared images.

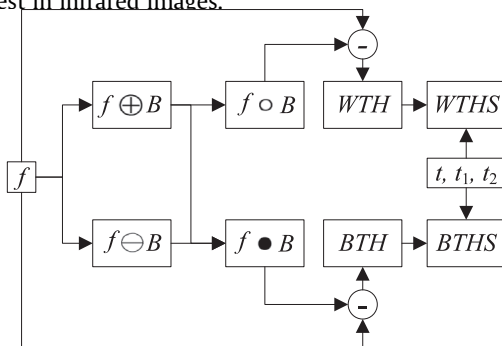


Fig. 1. Calculation of the top-hat selection transforms.

Also, in visual images, the regions of interest which are different from the surrounding regions could be extracted using the similar way. Therefore, these extracted regions of interest which are important image features could be used for image fusion.

Top-hat transform is an important operation in mathematical morphology and has been widely used in different image processing applications [21–24]. However, because of the detail smoothing of opening or closing operation in top-hat transform, top-hat transform could not well utilize the difference information between the regions of interest and the surrounding background regions [21–24]. So, the performance of top-hat transform for image processing is not good in some cases. To well utilize the difference information between the regions of interest and the surrounding background regions, the top-hat selection transform constructed through selectively output the difference information between the regions of interest and the surrounding background regions has been given [24]. Through designing the selection rules in the top-hat selection transform would result in effective tools for different applications.

In this paper, the regions of interest usually have different gray values comparing with the surrounding background regions. So, it would be easy to discriminate the regions of interest through importing a value into the selection rules in top-hat selection transform. The important image features in the original infrared and visual images are usually different from the surrounding background regions. These features may be well extracted through using a value which could discriminate the important features. This is useful for infrared and visual image fusion. Therefore, the strategy of simply specifying the top-hat selection transform for the purpose of infrared and visual image fusion in the expressions (10) and (11) is simple but reasonable.

Expressions (10) and (11) require a minimum contrast in the infrared images. The purpose of this paper is to combine the important regions of interest in the original infrared and visual images into the final fusion image. Because the regions of interest are usually different from the surrounding regions, which ensures that the needed contrast exists. So, the definition of expressions (10) and (11) is reasonable.

If the contrast of the infrared image is very low, the value nL used in the expressions (10) and (11) could be valued as a small value, which will also effectively extract the needed image features. Moreover, the experimental results, including the qualitative and quantitative comparisons with some recent and effective algorithms, verified the effective performance of the proposed algorithm. Therefore, the proposed algorithm is an effective way for infrared and visual image fusion.

For the case of noisy images, because the noises in image may be different from the surrounding regions, the proposed algorithm may also extract the noises and combine them into the final fusion image, which would affect the performance of the proposed algorithm. Fortunately, the images used in the application of this paper do not contain heavy noises. And, the experimental results verified that the proposed algorithm performed well on these images.

3.1.2. Multiscale top-hat selection transform based region extraction

Top-hat selection transform extracts image regions with size corresponding to the size of the used structuring element. Usually, the useful image regions have different sizes and exist at different scales. To extract all the useful image regions for image fusion, multiscale structuring elements with different sizes should be used [24].

Suppose n scales of structuring elements, $B_1, B_2, \dots, B_n$, which have the same shape and increasing sizes, should be used. $B_i = \underbrace{B_1 \oplus B_1 \dots \oplus B_1}_{\text{dilation } i \text{ times}}, 1 \leq i \leq n$.

Based on the multiscale structuring elements, the bright and dim image regions in image f at each scale could be extracted as follows,

$$WTHS_i(x, y) = \begin{cases} f(x, y) - f \circ B_i(x, y), & f(x, y) - f \circ B_i(x, y) \geq nL \\ 0, & \text{else,} \end{cases} \quad (12)$$

$$BTHS_i(x, y) = \begin{cases} f \bullet B_i(x, y) - f(x, y), & f \bullet B_i(x, y) - f(x, y) \geq nL \\ 0, & \text{else} \end{cases} \quad (13)$$

$WTHS_i$ and $BTHS_i$ are the extracted bright and dim image regions corresponding to the scale i using structuring element B_i .

3.1.3. Multiscale region extraction for image fusion

Let f_{IR} and f_{VI} represent the original infrared and visual images. Based on the multiscale top-hat selection transform, the extracted bright image regions of f_{IR} at scale i could be expressed as follows,

$$[WTHS_i(f_{IR})](x, y) = \begin{cases} f_{IR}(x, y) - f_{IR} \circ B_i(x, y), & f_{IR}(x, y) - f_{IR} \circ B_i(x, y) \geq nIR \\ 0, & \text{else.} \end{cases} \quad (14)$$

nIR is the parameter used by top-hat selection transform to extract the bright image regions in infrared image.

Also, the bright image regions of f_{VI} at scale i could be expressed as follows,

$$[WTHS_i(f_{VI})](x, y) = \begin{cases} f_{VI}(x, y) - f_{VI} \circ B_i(x, y), & f_{VI}(x, y) - f_{VI} \circ B_i(x, y) \geq nVI \\ 0, & \text{else.} \end{cases} \quad (15)$$

nVI is the parameter used by top-hat selection transform to extract the bright image regions in visual image.

Similarly, the extract dim image regions of f_{IR} and f_{VI} at scale i could be expressed as follows,

$$[BTHS_i(f_{IR})](x, y) = \begin{cases} f_{IR} \bullet B_i(x, y) - f_{IR}(x, y), & \text{if } f_{IR} \bullet B_i(x, y) - f_{IR}(x, y) \geq nIR \\ 0, & \text{else,} \end{cases} \quad (16)$$

$$[BTHS_i(f_{VI})](x, y) = \begin{cases} f_{VI} \bullet B_i(x, y) - f_{VI}(x, y), & \text{if } f_{VI} \bullet B_i(x, y) - f_{VI}(x, y) \geq nVI \\ 0, & \text{else.} \end{cases} \quad (17)$$

3.1.4. Specifying of nIR and nVI

nIR and nVI are important parameters used in top-hat selection transform to extract the regions of interest, which is mainly valued based on the gray value difference between the regions of interest and surrounding regions.

In infrared image, the regions of interest are bright or dim regions. And, the important regions of interest in infrared image are usually different from the surrounding regions. If the regions of interest are very bright or dim, nIR could be a large value to suppress the most of the other background regions. This would be useful to produce a clear fusion image with useful image information from the original images. Otherwise, nIR should be a small value to maintain the important image regions in the final fusion image. Correspondingly, if the regions of interest are very bright and dim, the gray value varying in image would be large, which results in a large standard deviation value of the gray values of the image. And, if the regions of interest are not very bright and dim, the standard deviation value of the gray values of the image will be small. Therefore, nIR could be valued following the standard deviation value of the gray values of image. Let σ_{IR} represent the standard deviation value of the gray values of the original infrared image. nIR is valued as follows in this paper.

$$nIR = 0.4 \times \sigma_{IR}. \quad (18)$$

nIR is smaller than σ_{IR} , which will ensure that nIR is smaller than the gray value difference between the regions of interest and the background regions. Then, the regions of interest could be extracted by top-hat selection transform. Also, nIR is changed following σ_{IR} , which may ensure that all the regions of interest could be extracted and the surrounding regions are suppressed. This would be useful for constructing the clear and effective fusion result.

Visual image usually contains many image details and the regions of interest may be not very different from the surrounding regions. Then, nVI should be a small value, so that the important image regions will be well maintained in the result of top-hat selection transform. nVI is valued as follows in this paper.

$$nVI = 0.2 \times \sigma_{VI}. \quad (19)$$

σ_{VI} is the standard deviation value of the gray values of the original visual image. nVI has the similar property as the nIR for region extraction by using top-hat selection transform.

Actually, nIR and nIV are mainly used to suppress the possible backgrounds in the extracted features by top-hat transform at each scale following the varying of the gray values in an image, so that the regions of interest could be well extracted. Usually, because the extracted image features by top-hat transform are mainly useful image regions, the possibly suppressed backgrounds do not have large gray values in the result of top-hat transform. Thus, they could be easily suppressed by using a small and reasonable threshold value. nIR and nIV are defined based on the variance of the gray values of the original image with a small weight, which are usually small values and reasonable for discriminating the regions of interest and surrounding backgrounds. So, using the same values of nIR and nIV for all the scales is also effective. And, experimental results on all the images show the good performance of the proposed algorithm using this definition, which verified that this definition could be used for all the images and did not need to be changed from one image to another image.

Also, using the same values of nIR and nIV for all the scales will simplify the implementation of the proposed algorithm, which makes the proposed algorithm more applicable.

According to the definition, if the contrast of the original image is good, the extracted regions of interest may be very different from the surrounding regions. In this case, the standard deviation value of the original image would be large and thus leading to large nIR and nVI , which would well extract the regions of interest and suppress the affect of other regions. Conversely, if the contrast of the original image is low, the extracted regions of interest may be not very different from the surrounding regions. In this case, the standard deviation value of the original image would be small and thus leading to small nIR and nVI , which would also benefit the extracting of the regions of interest in images with low contrast. Therefore, setting the values of nIR and nVI in this way is reasonable. So, the proposed algorithm could be well used for images with different contrasts, not only the simple cases.

Moreover, experimental results on different types of infrared and visual images have verified that the proposed algorithm was effective and performed better than some other algorithms.

It should be pointed out that, defining the more appropriate values of nIR and nIV for different scales may be one way of further improving the performance of the proposed algorithm. This would be further addressed in our future work.

And, the proposed algorithm performs the fusion of infrared and visual images through extracting the regions of interest which are different from the surrounding regions. Because the noises in image are usually different from the surrounding regions, the original infrared and visual images with noises may affect the performance of the proposed algorithm. However, usually, the original images may not contain very heavy noises. And, the image smoothing techniques could be used to smooth the noises before the

image fusion. This pre-processing will improve the performance of image fusion for images with noises.

Therefore, because of nIR and nVI , top-hat selection transform could effectively extract the regions of interest and suppress the surrounding background regions, which will result in a clear fusion result image containing rich image details.

3.2. Image fusion

3.2.1. Constructing the final fusion regions

The extracted image regions at each scale are bright regions in the result of top-hat selection transform. Then, the fusion regions of the infrared and visual images at each scale should be the pixel-wise maximum of the extracted infrared and visual regions at each scale. So, the bright ($WTHS_i$) and dim ($BTHS_i$) fusion regions of scale i could be expressed as follows,

$$WTHS_i = \max\{WTHS_i(f_{IR}), WTHS_i(f_{VI})\}, \quad (20)$$

$$BTHS_i = \max\{BTHS_i(f_{IR}), BTHS_i(f_{VI})\}. \quad (21)$$

The extracted image regions in the result of top-hat selection transform have large gray values. Thus, the extracted image region at the scale should have larger gray values than the same region at other scales. Therefore, the final bright (RB) and dim (RD) fusion regions of all the scales could be obtained by applying the pixel-wise maximum operation on the extracted multiscale bright and dim image regions at all the scales as follows, respectively.

$$RB = \max_i\{WTHS_i\}, \quad (22)$$

$$RD = \max_i\{BTHS_i\}. \quad (23)$$

These constructed final fusion regions could be used to calculate the final fusion image.

To clearly show the final fusion regions and the extracted regions of interest at each scale, an example is shown in Figs. 2–4. The gray values of images in Figs. 3 and 4 have been scaled to an appropriate interval to obtain a good display effect to facilitate the visual observation of the readers. In Fig. 2a and b are the original infrared and visual images. Because there is no need to show all the scales, we use one low scale and one high scale which are the second and sixth scales to show the effect of extracting image features at each scale. The extracted bright and dim and the fusion regions of the second and sixth scales from the original infrared and visual images in Fig. 2 are shown in Fig. 3a and b are the extracted bright image regions of the second scale from the original infrared and visual images, denoted by $WTHS_2(f_{IR})$ and $WTHS_2(f_{VI})$, respectively. (c) is the bright fusion regions of the second scale $WTHS_2$. (d and e) are the extracted dim image regions of the second scale from the original infrared and visual images, denoted by $BTHS_2(f_{IR})$ and $BTHS_2(f_{VI})$, respectively. (f) is the dim fusion regions of the second scale $BTHS_2$. (g and h) are the extracted bright image regions of the sixth scale

from the original infrared and visual images, denoted by $WTHS_6(f_{IR})$ and $WTHS_6(f_{VI})$, respectively. (i) is the bright fusion regions of the sixth scale $WTHS_6$. (j and k) are the extracted dim image regions of the sixth scale from the original infrared and visual images, denoted by $BTHS_6(f_{IR})$ and $BTHS_6(f_{VI})$, respectively. (l) is the dim fusion regions of the sixth scale $BTHS_6$.

Fig. 3 shows that, the extracted bright or dim image regions of each scale are indeed the important bright or dim image regions in the original images. And, the extracted image regions are important image information in the original images, which are the useful information and should be combined into the final fusion image. Fig. 3 also indicates that, the extracted bright and dim image regions at each scale have large gray values. So, combining the extracted image regions at each scale using the pixel-wise maximum operation is reasonable. For example, in Fig. 3a and b, the extracted bright image regions of the second scale from the original infrared and visual images are brighter than other regions. Thus, after the pixel-wise maximum operation, the bright fusion regions of the second scale well combines these extracted bright image regions from the original infrared and visual images, which is shown in Fig. 3c. Moreover, the extracted image regions at the low scale (Fig. 3a, b, d and e) usually represent the important image regions with small size in images, and the extracted image regions at the high scale (Fig. 3g, h, j and k) usually represent the important regions in images. So, the algorithm well extracts the important image regions at different scales, which contain the useful image information of all the scales. Therefore, combining these extracted image regions into the final fusion image will results in an effective fusion image with the extracted image regions from the original infrared and visual images.

Fig. 3 also shows that the extracted image features (especially the regions of interest) are different in the low and high scales. In the second scale which is a low scale, the extracted image features are mainly the image details and the regions of interest with small size. And, in the sixth scale which is a high scale, the extracted image features are mainly the regions of interest with large size. Therefore, to extract all the image features, especially the regions of interest, at different scales for the effective infrared and visual image fusion, multiscales of top-hat selection transform should be used.

The final bright and dim fusion regions of all the scales, denoted by RB and RD , are shown in Fig. 4a and b, respectively. Because the fusion image regions of each scale have large gray values, the final fusion image regions could be easily obtained by using the pixel-wise maximum operation on all the fusion image regions of all the scales as expressions (22) and (23) shown. Fig. 4 shows that, the extracted bright and dim image regions from the original infrared and visual images are well combined together. And, these final fusion regions are indeed the important regions in the original infrared and visual images. This would be useful for obtaining an effective fusion image.

Because there is no need to show all the scales, we use one low and one high scales which are the second and sixth scales to show the effect of extracting image features at each scale. And, the gray values of the extracted image regions are large at each scale comparing with other scales, the pixel-wise maximum operation is used to combine the extracted image features at all the scales as expressions (22) and (23) shown.

3.2.2. Adaptive weight strategy based image fusion

Combining these extracted image regions into a base image which contains the basic information of the original images will produce an effective fusion result. And, the pixel-wise averaging of the original images could be used as a base image containing the basic information of the original images as follows,



Fig. 2. The original infrared and visual images.

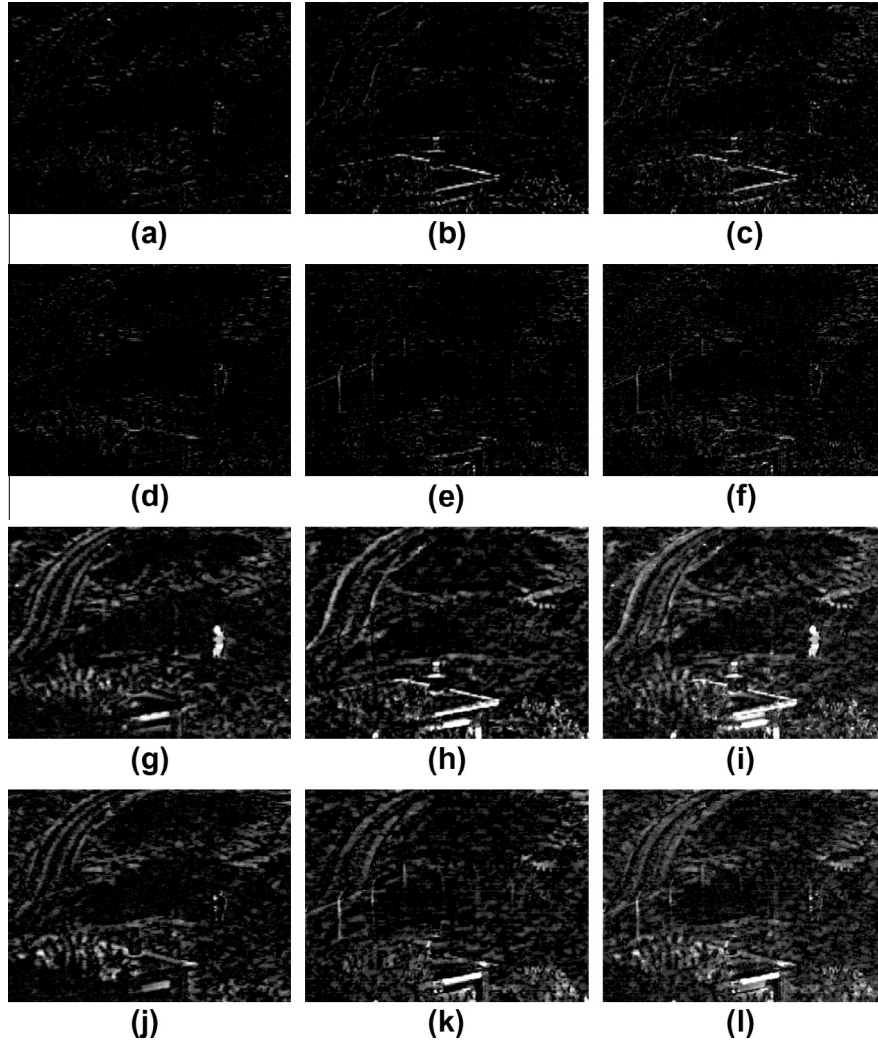


Fig. 3. The extracted bright and dim and the fusion regions of the second and six scales from the original infrared and visual images.

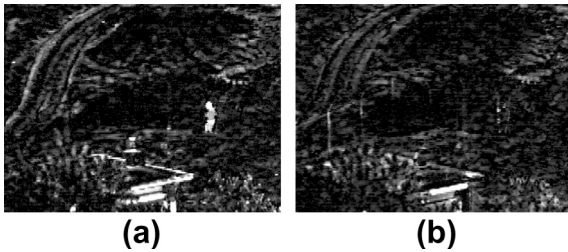


Fig. 4. The final bright and dim fusion regions of all the scales.

$$RA(x, y) = \frac{f_{IR}(x, y) + f_{VI}(x, y)}{2}. \quad (24)$$

RB and RD are the extracted final bright and dim fusion regions which represent the extracted regions of interest by the multiscale top-hat selection transform for image fusion. The important image regions in infrared image usually have good contrast comparing with the surrounding background regions. So, importing the extracted bright and dim image regions into the base image through contrast enlargement may result in a fusion result image which has good visual effect and is clear.

Moreover, bright or dim image regions with larger gray values contain more important image features which should be remained

in the final result image more. So, the mean gray value of RB and RD , denoted by pb and pd , are used as the adaptive weights in the fusion result. Thus, the image fusion could be expressed as follows,

$$F_F = RA + RB \times pb - RD \times pd, \quad (25)$$

where

$$pb = \text{mean}_{x,y}(RB), \quad (26)$$

$$pd = \text{mean}_{x,y}(RD). \quad (27)$$

In f_F , the final bright fusion regions are added on and the final dim fusion regions are subtracted from the base image. Only using the multiscale technique in the procedure of the region extraction may lead to the loss of contrast information in the fused image. However, the strategy of adding the final bright fusion regions on and subtracting the final dim fusion regions from the base image will enhance the contrast of the image. So, the contrast of the fusion image would be good. And, the weights pb and pd ensure that, the bright or dim image regions which contains more image regions are adaptively remained in the final result image more, which will further enhance the contrast of the final fusion result. Therefore, the final fusion result will be effective and the image details will be clear.

3.3. Structuring element selection

Structuring element is one important parameter in mathematical morphology based algorithm. And, the shape and size of structuring element should be determined.

The widely used shapes are rectangle, square, rhombus and circle. However, because the circle shape does not have sharp corners and is benefit for suppressing the block effect, the circle shape has been widely used in different applications. So, in this paper, the shape of the structuring element is circle. Other types of shapes, such as rectangle, square and rhombus could be also used for the reason of easy implementation. The experimental results also verified that, using circle shape was effective for all the used image sets in this paper.

In this paper, the size of the structuring element is determined by the scale number. A large scale number will extract more image regions of more scales using larger structuring element, which may improves the performance of the proposed algorithm. But, usually, there is no need to use a very large scale number. And, we have tested the proposed algorithm on image sets from different applications. The results show that, using 6 or 7 as the scale number would be enough for all the images. So, in this paper, $n = 6$. The radius of the circle structuring element at each scale is the number of each scale.

3.4. Implementation

The implementation of the proposed algorithm is illustrated in Fig. 5. First, the multiscale bright and dim image regions are extracted from the original infrared and visual images by using the multiscale top-hat selection transform, respectively. Second, the bright and dim fusion regions of each scale are calculated through the pixel-wise maximum operation. Third, the final bright and dim fusion regions of all the scales are produced based on the bright and dim fusion regions of each scale. Finally, the final fusion result is obtained by importing the final bright and dim fusion regions into the base image using a weight strategy.

4. Experimental results

Five image sets have been used in the experiment to verify the performance of the proposed algorithm. And, to do the comparison, the widely used wavelet pyramid algorithm (WP) [7,8], shift invariant discrete wavelet transform based algorithm (SIDWT) [10], Laplacian pyramid algorithm (LP) [14], morphological pyramid algorithm (MP) [21], multiscale morphology based algorithm (MSM) [5], center-surround top-hat transform based algorithm (CSTHT) [23] and the proposed algorithm have been applied on these image sets. The WP, SIDWT and LP are effectively and widely used multiscale technique based algorithms. The MSM and CSTHT are multiscale morphology based algorithms through designing morphological operators. MP is a multiscale technique based algorithm by combining the superiorities of pyramid technique and morphological operators. The proposed algorithm is also a multiscale technique based morphological algorithm and intends to be an effective image fusion algorithm. So, the WP, SIDWT, LP, MP, MSM and CSTHT are used to do the comparison in this paper.

4.1. Qualitative comparison experiment

Figs. 6–9 list some comparison examples on images from the five image sets. In these figures a and b are the original infrared and visual images. (c) is the fusion result of the wavelet pyramid algorithm (WP). (d) is the fusion result of the shift invariant discrete wavelet transform based algorithm (SIDWT). (e) is the fusion

result of the Laplacian pyramid algorithm (LP). (f) is the fusion result of the morphological pyramid algorithm (MP). (g) is the fusion result of the multiscale morphology based algorithm (MSM). (h) is the fusion result of the center-surround top-hat transform based algorithm (CSTHT). (i) is the fusion result of the proposed algorithm.

Fig. 6 is a comparison example on images from the “OctecWS” image set. In Fig. 6c, because the people target regions could not be easily recognized and the difference between the bright regions and dim regions is not good, the contrast of the result of WP (Fig. 6c) is not good. Also, the image details in this result are not clear, which result in an un-clear image. MP produces many artifacts (Fig. 6f), which destroys some useful information of the original images. The contrast of the result of CSTHT (Fig. 6h) is good, but the image is not clear because some image details are smoothed. Comparing with LP, SIDWT and the proposed algorithm, MSM still smoothes some image details (Fig. 6g) which produce a not clear result image. Because the image details in the results of SIDWT (Fig. 6d), LP (Fig. 6e) and the proposed algorithm (Fig. 6i) are rich and the difference between the bright and dim image regions are good in these result images, SIDWT (Fig. 6d), LP (Fig. 6e) and the proposed algorithm (Fig. 6i) not only make the image details clear, but also get fusion results with good contrast.

The proposed algorithm produces some artifacts in the text areas. The reason is these artifacts are extracted as the regions of interest in the original visual image, and thus they are maintained in the final fusion image. Actually, the text regions are not important information in the original images. So, even there are some artifacts, the further image analysis will not be affected. In another way, because the text areas are regions added on the image manually, they are different from the surrounding regions. Thus, effectively extracting these regions by the proposed algorithm also verifies the good performance of the proposed algorithm for infrared and visual image fusion through extracting the regions of interest. Therefore, the proposed algorithm performs well for infrared and visual image fusion. And, in Fig. 6, the result of Laplacian pyramid algorithm introduces some artifacts around the tree area, which may affect the analysis of the fusion image. The proposed algorithm well maintains the image details comparing with other algorithms and does not introduce many artifacts except the unimportant text areas. Therefore, the performance of the proposed algorithm is effective.

Fig. 7 is a comparison example on images from the “UNcamp” image set. The results show that, all the algorithms could combine the useful image information of the original images and achieve the purpose of image fusion. Because some image details are smoothed and the difference between the bright and dim regions is not good, the result of WP (Fig. 7c) is not clear and the contrast is not good. The contrast of CSTHT (Fig. 7h) is good, but the result is not clear because some image details are smoothed. The results of LP (Fig. 7e) and MP (Fig. 7f) are clearer than CSTHT and the contrast is good, but the result images are not clear. SIDWT (Fig. 7d) and MSM (Fig. 7g) smooth some image details, especially in the regions of road. Because the regions of interest are well extracted and combined into the final fusion result, the result of the proposed algorithm (Fig. 7i) is clearer than other algorithms. Although the contrast of the people target region may be a little lower than other algorithms, the contrast is still good and the image details of the whole image are very clear. For example, the fence in the image is clearer than other results, which is because the regions of interest are well combined in the final fusion image. So, overall, the performance of the proposed algorithm is effective.

Fig. 8 is a comparison example on images from the “Navi” image set. The original infrared image gives some important information, such as the road regions. And, the original visual image contains many visual image details, such as the tree regions and

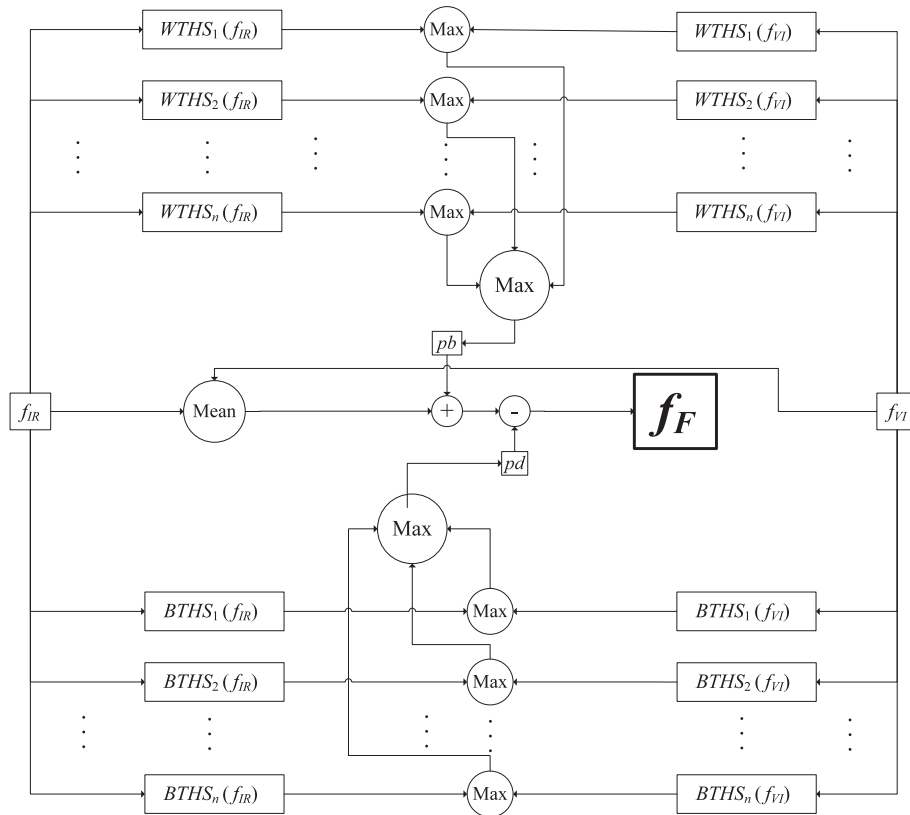


Fig. 5. Implementation of the proposed algorithm.

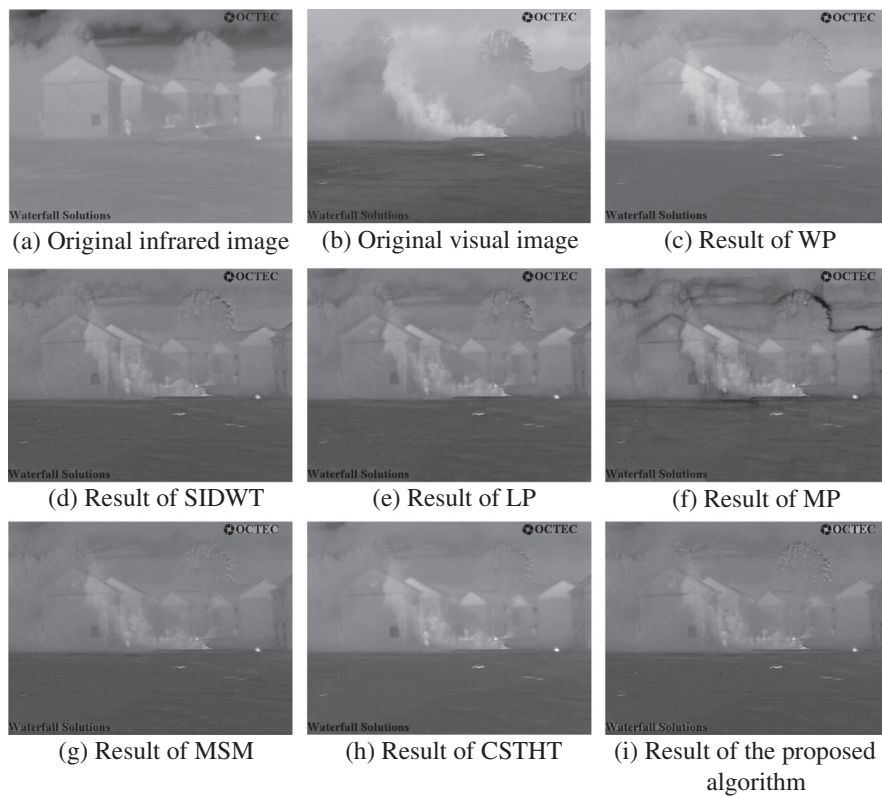


Fig. 6. An example on OctecWS images.

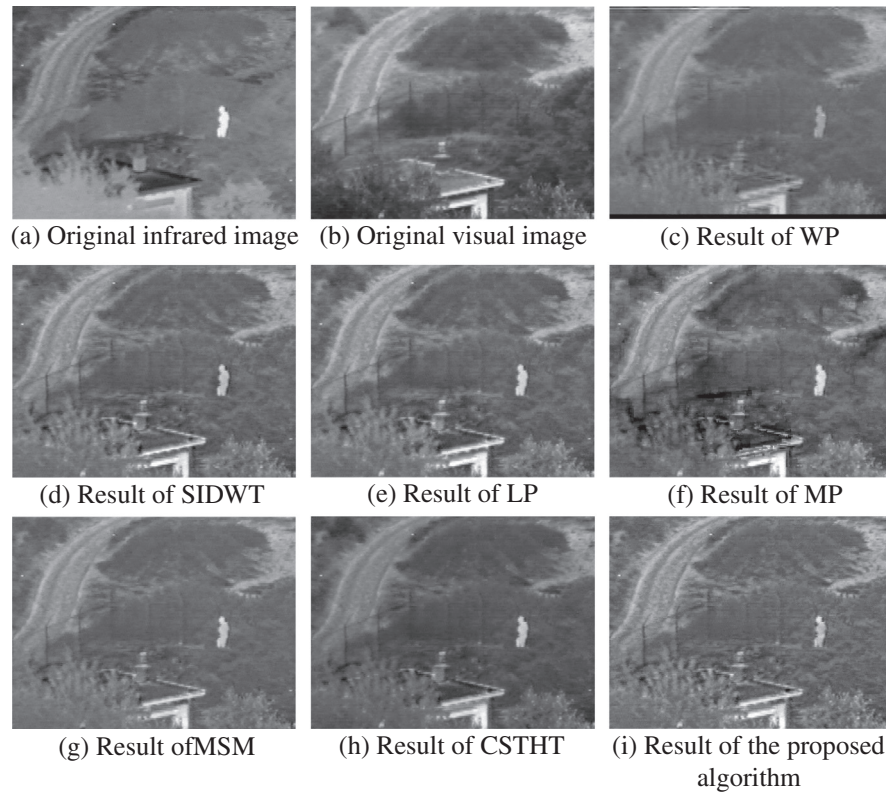


Fig. 7. An example on UNcamp images.

some other regions. The result images show that, all the algorithms could combine these useful image regions. Because some image details are smoothed by WP and the difference between the bright and dim regions is not good, the result of WP (Fig. 8c) is not clear and the contrast is not good. The contrast of the results of the rest algorithms is good, because the difference between the bright and dim regions in these result images is good. But, CSTHT (Fig. 8h) and MSM (Fig. 8g) smooth many image details, which results in not clear images. MP (Fig. 8f) produces many artifacts, which will affect the application of the fusion image. LP (Fig. 8e) and SIDWT (Fig. 8d) perform better than other algorithms. But, comparing with the result of the proposed algorithm (Fig. 8i), the contrast of the result of the proposed algorithm is good and there are more useful regions of interest in the result of the proposed algorithm because the proposed algorithm could well extract the regions of interest through the multiscale top-hat selection transform.

Fig. 9 is a comparison example on images from the “Trees” image set. The original images contain many tree details and the “people” region is the important region of interest. The result of WP (Fig. 9c) is not clear because some image details are smoothed. MP produces many artifacts in the result image (Fig. 9f). MSM smooths some image details (Fig. 9g). Although the result of the CSTHT (Fig. 9h) has a good contrast, the result is not clear because some image details are smoothed. The results of SIDWT (Fig. 9d) and LP (Fig. 9e) are better than MSM. But, comparing with the result of the proposed algorithm, the results of LP and SIDWT are not clear. Because the proposed algorithm could well extract the regions of interest, the result of the proposed algorithm (Fig. 9i) is clear and the important image regions, such as “people” region, are well maintained. So, the proposed algorithm performs better than some other algorithms.

In Figs. 8 and 9, because one of the original images is dark, some algorithms remain more information of the dark image in the final fusion image. Although the contrast of these fusion results looks good, the information of another original image is not well pre-

served. So, the image details are not clear in these results. The proposed algorithm well combines the useful information of all the original images, which results in a clear and effective fusion result.

To give an overall qualitative comparison, Table 1 shows some aspects of the performances of these algorithms, including the detailed features, clarity, contrast and artifacts reducing. These results are obtained based on the overall quality comparisons on all the image sets performed by the researchers in our laboratory. “Some” means the fusion result contains some corresponding features but the performance is not good. “Acceptable” means the fusion result is acceptable corresponding to the features and the performance is better than “Some”. “Good” means the performance is the best among these results.

Table 1 shows that, the clarity of SIDWT, LP and the proposed algorithm is better than other algorithms. And, the contrast of SIDWT, LP and the proposed algorithm is good. Also, SIDWT, LP and the proposed algorithm do not produce many artifacts. So, the performances of SIDWT, LP and the proposed algorithm are better than other algorithms. However, in the visual results in Figs. 6–9, the contained rich details in the results of the proposed algorithm are the most among these algorithms. Thus, the proposed algorithm performs well for infrared and visual image fusion, and the performance is better than other algorithms.

All of these experimental results show that, because the proposed algorithm could well extract the regions of interest of the original images by using the multiscale top-hat selection transform, the fusion result of the proposed algorithm is effective and clear. Then, the result image could be well used for different applications, such as target detection, object recognition, and image navigation.

4.2. Quantitative comparison experiment

Effective infrared and visual image fusion algorithm should well extract the important regions of interest and combine them into

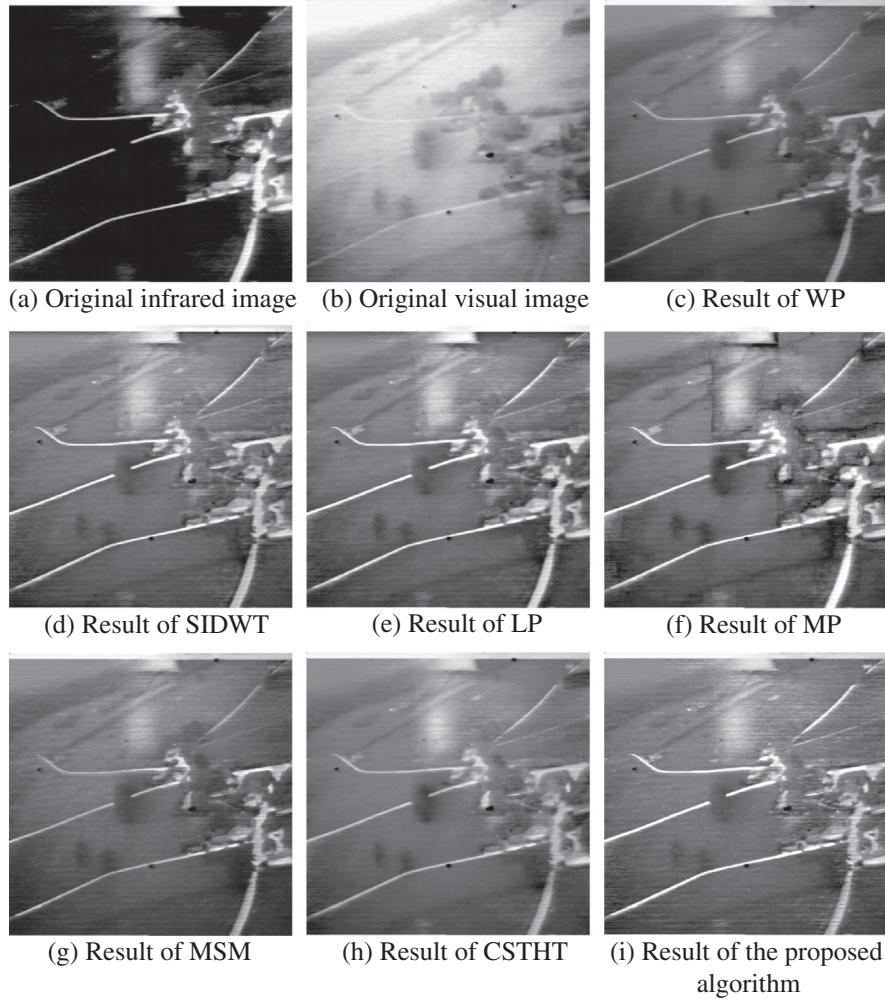


Fig. 8. An example on Navi images.

the final fusion result image. Thus, the final result image should be clear and have good contrast. Moreover, for the purpose of fusion, the fusion result image should effectively import the original information of the original infrared and visual images into the final fusion image. And, the fusion result image should contain rich image details.

To well evaluate the quantitative performance of the proposed algorithm and do the comparison, three measures, which are the spatial frequency (SF) [26], mean gradient (MG) [27] and image quality based measure (Q) [28], are used in this paper. Other measures could be also used in this paper for the performance evaluation [6,11,27–31].

Suppose the fusion image f_F has size $M \times N$. The calculation of SF is as follows,

$$SF = \sqrt{RF^2 + CF^2}, \quad (28)$$

where

$$RF = \sqrt{\frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N [f_F(x, y) - f_F(x-1, y)]^2}, \quad (29)$$

$$CF = \sqrt{\frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N [f_F(x, y) - f_F(x, y-1)]^2}. \quad (30)$$

The calculation of MG is as follows,

$$MG = \frac{1}{(M-1) \times (N-1)} \times \sum_{x=1}^{M-1} \sum_{y=1}^{N-1} \sqrt{(f_F(x, y) - f_F(x-1, y))^2 + (f_F(x, y) - f_F(x, y-1))^2} / 2 \quad (31)$$

Suppose the fusion image is f_F , and the corresponding original images are f_{IR} and f_{VI} . The calculation of Q is as follows,

$$Q(f_{IR}, f_{VI}, f_F) = \frac{1}{W} \sum_{w \in W} (\lambda(w) Q_0(f_{IR}, f_F|w) + (1 - \lambda(w)) Q_0(f_{VI}, f_F|w)), \quad (32)$$

where

$$\lambda(w) = \frac{s(f_{IR}|w)}{s(f_{IR}|w) + s(f_{VI}|w)}. \quad (33)$$

$Q_0(f_{IR}, f_F|w)$ is the computed result using the pixel values of f_{IR} and f_F in the small window w . Details about the selection of w are in [28]. W is the family of w . $s(f_{IR}|w)$ and $s(f_{VI}|w)$ could be valued as the variance of the pixels in w of f_{IR} and f_{VI} , respectively. Suppose a and b are the valued sequences of the pixel values in the w of f_{IR} and f_F , then

$$Q_0(f_{IR}, f_F|w) = \frac{\sigma_{ab}}{\sigma_a \sigma_b} \times \frac{2\bar{a}\bar{b}}{\bar{a}^2 + \bar{b}^2} \times \frac{2\sigma_a \sigma_b}{\sigma_a^2 + \sigma_b^2}, \quad (34)$$

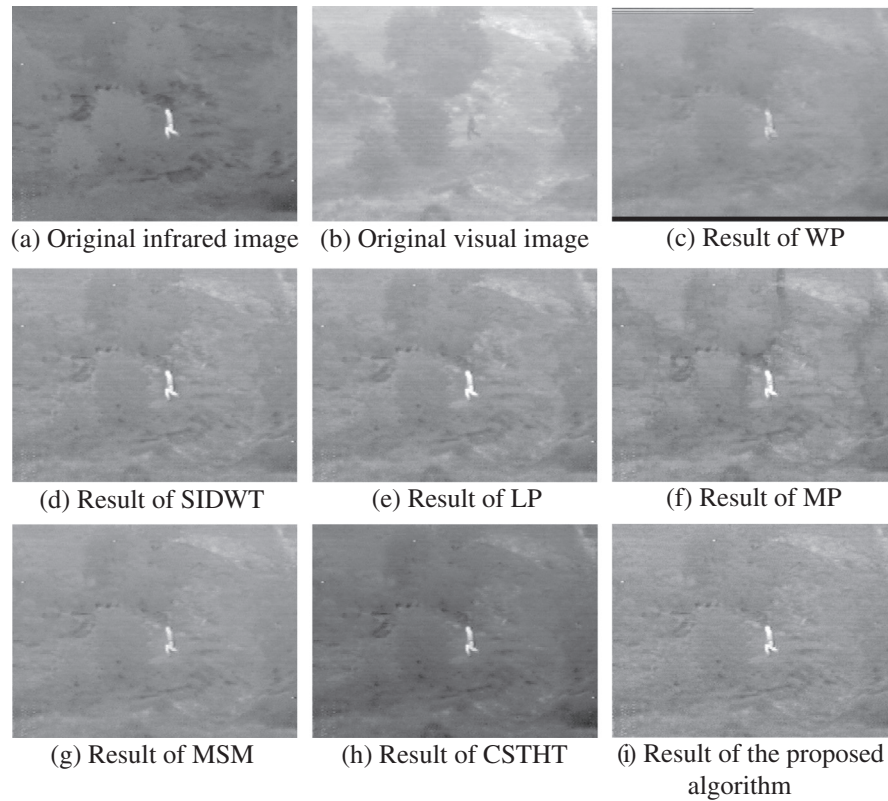


Fig. 9. An example on Trees images.

Table 1

Overall qualitative comparison.

	WP	SIDWT	LP	MP	MSM	CSTHT	The proposed algorithm
Detailed features	Some	Acceptable	Acceptable	Some	Acceptable	Some	Good
Clarity	Some	Good	Good	Some	Acceptable	Acceptable	Good
Contrast	Some	Good	Good	Good	Good	Good	Good
Artifacts reducing	Good	Good	Good	Some	Good	Good	Good

where, $\bar{a}$ and $\bar{b}$ are the mean value of the pixel values in a and b , respectively. σ_a and σ_b are the variance of the pixel values in a and b , respectively. σ_{ab} is the covariance of the pixel values in a and b . $Q_0(f_V, f_I|w)$ could be calculated using the similarly way.

The measure SF is one well used measure in image fusion [26]. SF is defined based on the contained spatial image details and contrast. Therefore, SF is an appropriate measure for quantitative comparison of the performances of different algorithms.

Moreover, using only one measure may be not enough for performance comparison because different measures may give different evaluation results [6,11,27–31]. So, to compare the performances of the image fusion algorithms using more measures, the well used gradient based measure which is mean gradient (MG) [27] and image quality based measure (Q) [28] are used in this paper. MG is a gradient based measure which has been well used in image fusion for measuring the clarity and contrast of image [27]. Q is defined using the quality of image and based on the performance of algorithm for importing the original information of the original infrared and visual images into the final fusion image. Big values of SF, MG and Q indicate good performance of the corresponding algorithm for infrared and visual image fusion.

Images from five image sets are processed by the proposed algorithm and the comparison algorithms. The final fusion results of the images produced by each algorithm are used to calculate the values of quantitative measures. And, the mean value of the

values of each measure corresponding to each algorithm on all the used images is shown in Figs. 10–12 to do the comparison, respectively. Figs. 10–12 list the quantitative comparisons using the measures SF, MG and Q, respectively. The image sets used in this paper to demonstrate the performance of the proposed algorithm are public data sets which do not contain much noise. Details of the comparison are shown below.

Fig. 10 is the quantitative comparison using the measure SF. The proposed algorithm gives a bigger value than other algorithms, which means the fusion results of the proposed algorithm contain more image details than other algorithms. Also, the contrast is good. All of these are because the proposed algorithm could well extract the regions of interest in the original infrared and visual images and effectively combine them into the final fusion image. So, the performance of the proposed algorithm for infrared and visual image fusion with emphasis on extracting regions of interest is better.

Fig. 11 is the quantitative comparison using the measure MG. The MG value of the proposed algorithm is bigger than other algorithms. This means, the fusion result image of the proposed algorithm is clear and the contrast is good. These are because the regions of interest in the original infrared and visual images are well extracted and used for image fusion. So, the fusion result of the proposed algorithm is better than other algorithms.

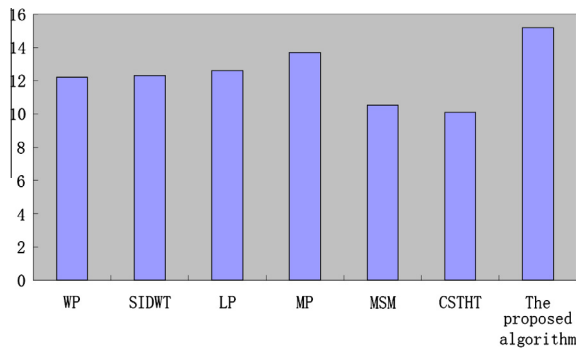


Fig. 10. Quantitative comparison using the measure SF.

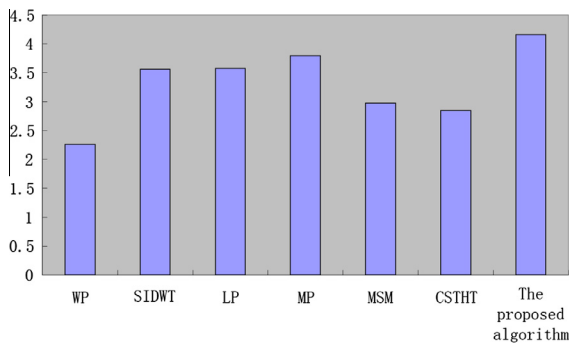


Fig. 11. Quantitative comparison using the measure MG.

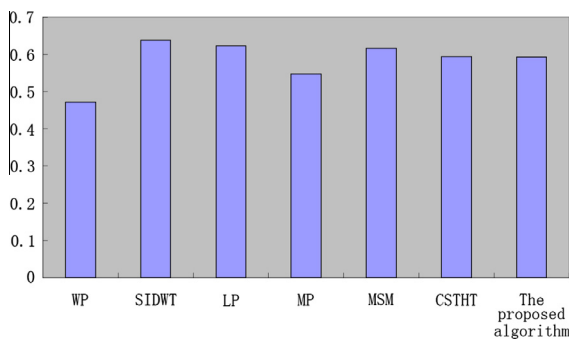


Fig. 12. Quantitative comparison using the measure Q.

Fig. 12 is the quantitative comparison using the measure Q. Although the Q value is not the largest one, the difference between the Q value of the proposed algorithm and the larger Q values (SIDWT, LP) are very small. This means, the proposed algorithm also performs well for importing the original information of the original infrared and visual images into the final fusion image.

Moreover, the SF and MG values of the proposed algorithm are larger than other algorithms. Therefore, the proposed algorithm could well extract and combine the regions of interest in the original images to achieve a clear fusion image with good contrast. And,

Table 2
Calculation time comparison of different algorithms (s).

WP	SIDWT	LP	MP	MSM	CSTHT	The proposed algorithm
0.574	0.733	0.082	1.311	0.923	25.450	3.874

the original image information is well maintained in the fusion result image of the proposed algorithm.

To compare the calculation time of different algorithms, the algorithms are performed on infrared and visual images with size 360×270 (CPU: Intel Pentium 4, 2.6 GHz; Memory: 512 MB). Each algorithm is performed several times. And, the mean time of each algorithm is used to do the comparison. The calculation time of each algorithm is shown in Table 2.

Table 2 shows that, the proposed algorithm is not very fast. Because the morphological operations have to be performed several times due to the multiscale morphological theory, the calculation of the multiscale morphological theory based algorithms (CSTHT, MP, MSM and the proposed algorithm) uses more time than other algorithms (WP, LP and SIDWT). And, because the calculation of the top-hat selection top-hat transform in the proposed algorithm is faster than the calculation of center surround top-hat transform in CSTHT, the calculation time of the proposed algorithm is very faster than CSTHT. But, because the top-hat selection transform is the modified version of the classical top-hat transform, and the comparison operation in the top-hat selection transform may occupy some calculation time, the calculation time of the proposed algorithm is longer than MSM. However, several methods [32,33] which have been well used to speed up the calculation of the morphological operations for real time applications could be used to decrease the calculation time of the proposed algorithm. Moreover, the proposed algorithm performs well on different types of infrared and visual images and the performance is better than some other algorithms. Therefore, the proposed algorithm is effective and useful for infrared and visual image fusion related applications, such as target detection, object recognition, and image navigation.

Also, to demonstrate the overall comparison of these algorithms based on the quantitative comparisons shown in Figs. 10–12 and Table 2, a new Table 3 is given. In Table 3, “Some” means the corresponding quantitative value is worse than most of other algorithms. “Acceptable” means the corresponding quantitative value is not very different from the best one. “Good” means the corresponding quantitative value is the best among these algorithms. Table 3 shows that, although the performances of the proposed algorithm on the Q value and the calculation time are not the best, the performances are acceptable and not very different from the best ones. More importantly, the performances of the proposed algorithm on the spatial frequency and mean gradient are the best among these algorithms. Therefore, overall, the performance of the proposed algorithm is effective. Table 3 further verifies the effective performance of the proposed algorithm for infrared and visual image fusion.

These experimental results verify the effective performance of the proposed algorithm for infrared and visual image fusion with emphasis on extracting the regions of interest of different images from different applications. So, the proposed algorithm may be

Table 3
Overall comparison based on the quantitative comparisons.

	WP	SIDWT	LP	MP	MSM	CSTHT	The proposed algorithm
Spatial frequency (Fig. 10)	Some	Some	Acceptable	Acceptable	Some	Some	Good
Mean gradient (Fig. 11)	Some	Some	Some	Acceptable	Some	Some	Good
Q value (Fig. 12)	Some	Good	Acceptable	Some	Acceptable	Acceptable	Acceptable
Calculation time (Table 2)	Good	Good	Good	Acceptable	Good	Some	Acceptable

well used in applications related to infrared and visual image fusion.

5. Summary and conclusions

Infrared image contains important regions of interest while visual image shows many useful image details. The useful image details are also important regions of interest contained in the visual image. Image fusion is an effective technique to combine the useful information of the original infrared and visual images. To be effective for infrared and visual image fusion, a multiscale top-hat selection transform based algorithm with emphasis on extracting the regions of interest is proposed in this paper.

Top-hat selection transform well differentiates the regions of interest from the background regions, which extracts the regions of interest in the infrared and visual images. Thus, though combining the extracted useful image information, the proposed algorithm could produce a clear result which contains more useful image information. Therefore, because of the effective extraction of the regions of interest and the reasonable combination of the extracted image regions, the proposed algorithm performs well for infrared and visual image fusion. Moreover, automatic selection of the parameters in the top-hat selection transform and the fusion weights improve the adaptability of the proposed algorithm for different types of images. Experimental results show that, the proposed algorithm performs well on different types of infrared and visual images. Therefore, the proposed algorithm could be used in applications related to infrared and visual image fusion, such as target detection, object recognition, and image navigation. However, if in the night-time scenes, visual images may not contain many useful image details, which will result in an unclear image fusion result.

Acknowledgements

The authors thank the anonymous reviewers and editor for the very constructive comments. The authors are grateful to Dr. Yan Li at Peking University, Beijing, China, for many helpful discussions and comments. This work is partly supported by the National Natural Science Foundation of China (Grant No. 61271023), open funding Project of State Key Laboratory of Virtual Reality Technology and Systems (Grant No. BUAA-VR-12KF-04), Fundamental Research Funds for the Central Universities (Grant No. YWF-11-03-Q-065), Innovation Foundation of AVIC (Grant No. CXY2010BH02) and Beijing Key Laboratory of Digital Media. The authors acknowledge the used original images available at <http://www.imagefusion.org>.

References

- [1] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, R. Arbiol, Multiresolution-based image fusion with additive wavelet decomposition, *IEEE Trans. Geosci. Remote Sens.* 37 (3) (1999) 1204–1211.
- [2] Xiangzhi Bai, Fugen Zhou, Bindang Xue, Infrared dim small target enhancement using toggle contrast operator, *Infrared Phys. Technol.* 55 (2012) 177–182.
- [3] C.A. Lieber, S. Urayama, N. Rahim, R. Tu, R. Saroufeem, B. Reubner, Multimodal near infrared spectral imaging as an exploratory tool for dysplastic esophageal lesion identification, *Opt. Express* 14 (2006) 2211–2219.
- [4] M. Leviner, M. Maltz, A new multi-spectral feature level image fusion method for human interpretation, *Infrared Phys. Technol.* 52 (2009) 79–88.
- [5] Susanta Mukhopadhyay, Bhabatosh Chanda, Fusion of 2D grayscale images using multiscale morphology, *Pattern Recognit.* 34 (2001) 1939–1949.
- [6] C.S. Xydeas, V. Petrović, Objective image fusion performance measure, *Electron. Lett.* 36 (4) (2000) 306–309.
- [7] Gonzalo. Pajares, Jesus Manuel de la Cruz, A wavelet-based image fusion tutorial, *Pattern Recognit.* 37 (2004) 1855–1872.
- [8] Krista Amolins, Yun Zhang, Peter Dare, Wavelet based image fusion techniques – an introduction, review and comparison, *ISPRS J. Photogram. Remote Sens.* 62 (2007) 249–263.
- [9] L. Chipman, T. Orr, L. Graham, Wavelets and image fusion, in: *Proceedings of SPIE* 2569. (1995) 208–219.
- [10] Styliani Ioannidou, Vassilia Karathanassi, Investigation of the dual-tree complex and shift-invariant discrete wavelet transforms on Quickbird image fusion, *IEEE Geosci. Remote Sens. Lett.* 4 (1) (2007) 166–170.
- [11] A. Toet, E.M. Franken, Perceptual evaluation of different image fusion schemes, *Displays* 24 (1) (2003) 25–37.
- [12] Filippo Nencini, Andrea Garzelli, Stefano Baronti, Luciano Alparone, Remote sensing image fusion using the curvelet transform, *Inf. Fusion* 8 (2007) 143–156.
- [13] Myungjin Choi, Rae Young Kim, Myeong-Ryong Nam, Hong Oh Kim, Fusion of multispectral and panchromatic satellite images using the curvelet transform, *IEEE Geosci. Remote Sens. Lett.* 2 (2) (2005) 136–140.
- [14] D.M. Bulanova, T.F. Burks, V. Alchanatis, Image fusion of visible and thermal images for fruit detection, *Biosyst. Eng.* 103 (2009) 12–22.
- [15] Peter J. Burt, Edward H. Adelson, The Laplacian pyramid as a compact image code, *IEEE Trans. Commun. COM-31* (4) (1983) 532–540.
- [16] Nikolaos Mitianoudis, Tania Stathaki, Pixel-based and region-based image fusion schemes using ICA bases, *Inf. Fusion* 8 (2007) 131–142.
- [17] Nedeljko Cvejic, David Bull, Nishan Canagarajah, Region-based multimodal image fusion using ICA bases, *IEEE Sens. J.* 7 (5) (2007) 743–751.
- [18] M. González-Audicana, J.L. Saleta, R.G. Catalán, R. García, Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition, *IEEE Trans. Geosci. Remote Sens.* 42 (6) (2004) 1291–1299.
- [19] Zhaobin Wang, Yide Ma, Jason Gu, Multi-focus image fusion using PCNN, *Pattern Recognit.* 43 (2010) 2003–2016.
- [20] Wei Huang, Zhongliang Jing, Multi-focus image fusion using pulse coupled neural network, *Pattern Recognit. Lett.* 28 (2007) 1123–1132.
- [21] P. Soille, *Morphological Image Analysis-Principle and Applications*, Springer, Germany, 2003.
- [22] Xiangzhi Bai, Fugen Zhou, Analysis of new top-hat transformation and the application for infrared dim small target detection, *Pattern Recognit.* 43 (2010) 2145–2156.
- [23] Xiangzhi Bai, Fugen Zhou, Bindang Xue, Fusion of infrared and visual images through region extraction by using multiscale center-surround top-hat transform, *Opt. Express* 19 (9) (2011) 8444–8457.
- [24] Xiangzhi Bai, Fugen Zhou, Bindang Xue, Noise suppressed image enhancement using multiscale top-hat selection transform through region extraction, *Appl. Opt.* 51 (3) (2011) 338–347.
- [25] Ishita De, Bhabatosh Chanda, Buddhajyoti Chattopadhyay, Enhancing effective depth-of-field by image fusion using mathematical morphology, *Image Vision Comput.* 24 (2006) 1278–1287.
- [26] V. Aslantas, R. Kurban, A comparison of criterion functions for fusion of multi-focus noisy images, *Opt. Commun.* 282 (2009) 3231–3242.
- [27] Pushkar Pradham, Nicolas H. Younan, Roger L. King, R. García, Concepts of image fusion in remote sensing applications, in: Tania Stathaki (Ed.), *Image Fusion: Algorithms and Applications*, Academic Press, Great Britain, 2008, pp. 391–428.
- [28] G. Piella, H. Heijmans, A new quality metric for image fusion, in: *Proceedings of International Conference on Image Processing*, Barcelona, Catalonia, Spain, 2003, pp. 173–176.
- [29] V. Petrović, C.S. Xydeas, Evaluation of image fusion performance with visible differences, *Lect. Notes Comput. Sci.* 3023 (2004) 380–391.
- [30] G. Qu, D. Zhang, P. Yan, Information measure for performance of image fusion, *Electron. Lett.* 38 (7) (2002) 313–315.
- [31] Z. Liu, E. Blasch, Z. Xue, R. Langanieri, W. Wu, Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: a comparative survey, *IEEE Trans. Pattern Anal. Machine Intell.* 34 (1) (2012) 94–109.
- [32] F. Gonzalez, O. Tubio, F. Tobajas, V. De Armas, R. Esper-Chain, R. Sarmiento, Morphological processor for real-time image applications, *Microelectron. J.* 33 (2002) 1115–1122.
- [33] Hochong Park, Roland T. Chin, Decomposition of arbitrarily shaped morphological structuring elements, *IEEE Trans. Pattern Anal. Machine Intell.* 17 (1) (1995) 2–15.